

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 18, 2003 07:51:53 ; Search time 2838 Seconds  
(without alignments)

5342.693 Million cell updates/sec

Title: US-09-857-581-66  
Perfect score: 2396  
Sequence: 1 MLELALGLVIALPFRHLP.....AHSLVCVPLARIGVASKLLS 521

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

0=/cgn2\_1/USPTO\_spool/US09857581/runat\_12052003\_084115\_11910/app\_query.fasta.1.7

11  
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS-human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09857581 @CGN\_1\_1724 @runat\_12052003\_084115\_11910 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LANGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6 -FCAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.heg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2270	94.7	1563	AF195812	AF195812 Pisum sat
2	2269	94.7	1774	AF195798	AF195798 Glycine m
3	2268	94.7	1566	AF195806	AF195806 Vigna rad
4	2268	94.7	1566	AF195807	AF195807 Vigna rad
5	2268	94.7	1566	AF195808	AF195808 Vigna rad
6	2268	94.7	1566	AF195810	AF195810 Trifolium
7	2267	94.6	1566	AF195809	AF195809 Vigna rad
8	2267	94.6	1722	AF135484	AF135484 Glycine m
9	2251	93.9	1824	AF022462	AF022462 Glycine m
10	2251	93.9	1824	AF195799	AF195799 Glycine m
11	2201.5	91.9	1902	AF195819	AF195819 Glycine m
12	2201.5	91.4	1800	AF195818	AF195818 Glycine m
13	2189	91.1	1501	AF195801	AF195801 Medicago
14	2180	91.0	1501	AF195800	AF195800 Medicago
15	2179	90.9	1501	AF195815	AF195815 Trifolium
16	2179	90.9	1501	AF195815	AF195815 Trifolium

11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*  
29: em.vi:\*  
30: em.heg.hum:\*  
31: em.heg.inv:\*  
32: em.heg.other:\*  
33: em.heg.mus:\*  
34: em.heg.pin:\*  
35: em.heg.tod:\*  
36: em.heg.mam:\*  
37: em.heg.vrt:\*  
38: em.sy:\*  
39: em.hego.hum:\*  
40: em.hego.mus:\*  
41: em.hego.other:\*

Journal	Submitted (18-Oct-1999)	Nutrition and Health, 80:402, Wilmington, DE 19880-0402, USA
FEATURES	Location/Qualifiers	
source	1. .1563	
gene	/db_xref="taxon:3888"	
	1. .>1563	
	/gene="1f61"	
	1. .>1563	
CDS	/gene="1f61"	
	/note="cytochrome P450"	
	/codon_start=1	
	/product="isoflavone synthase 1"	
	/protein_id="AAF3453.1"	
	/db_xref="GI:6979548"	
	/translation="MLELALGFLVALFLHLPPTPSAKSKARHLPPNPPSPKPLPR I GHILDKDLTHVALIDLSKKHGFLEISFGSMPTVAVSTPELFLKFLDAHEATSF TFOFSVRLTLDVNSVAMVPGPVYKRVKRLIMDLIMLTIVTELRLPTQOIKKFLP RMAASAQAKPLDYTELIELKNTNSTISMMAGABEIRIDAREYAKITPESYSLDPT WPLKTLKQYERKIDILINKKDPVEVERIKRREIVARRKNGEVBEASGVFLDTIT LBFADLEMIKITEKQIKGLVVDFFSAGTDSVAATEMLDELINNRVLQAKAESEE YGVVSKRLADEVDTONLPYIRAIYKETFRMHPPLPVVYKCTCECEINGYVPEGSA VLEFNVGVGDGVKDRPSEPRERPLETGAEEAGVLDLRQGHOLLPGSGRMGMC GVNLTSMATTLASLIQCFDLQVIGPQOGLKGDRAKVMEEBAGLTVRAHSLVCV PLARIQVASTLS"	
BASE COUNT	399 a	414 c
ORIGIN	379 g	371 t

## Percent Similarity

[illegible]

QY 101 Phe\*\*\*ThrArpPheGlnThrSerAla\*\*\*Arg\*\*\*LeuThrTyArgp\*\*\*\*\*ValAla 120  
 DB 301 TTCACGACCAAGTTCACAACTCTGTAAGACGCTCACTTACGACCTCTGAGGC 360  
 QY 121 \*\*\*\*\*Pro\*\*\*GlyProTyTrp\*\*\*PheValArgGlyLeuIleMetAsnAspLeu 140  
 DB 361 ATGCTTCATTGGACCTTACTGAAAGTTCTGAGAGAGCTCATATGACACCTTCTC 420  
 QY 141 AsnAlaThrThrValAsn\*\*\*LeuArgProLeuArgThrGlnGlnIleArgGly\*\*\*Leu 160  
 DB 421 AACGCAACACGCTCAAGAGCTCAGAGCTTGAAGAGCCAAACATCCGAGATTCTCT 480  
 QY 161 Arg\*\*\*MetAlaGln\*\*\*AlaGluAla\*\*\*LysProLeuAsp\*\*\*ThrGlnGluLeu 180  
 DB 481 AGGGTTATGGCCCAAGCGCAGAGGCCCAAGCCCTTGATGATCAGTACCGAGAGCTTCTC 540  
 QY 181 LysTrp\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGlyGluAlaGluGlnIleArgAsp 200  
 DB 541 AATATGACCAACGACCATCTCCATGATGATGCTCGGCGAGGCTGAGAGATCAGAGC 600  
 QY 201 IleAlaArgGluValLeuLysIle\*\*\*GlyGluTySerLeuThrAspPheIle\*\*\*Pro 220  
 DB 601 ATCGCTCCGAGGTCCTTACATCTTGGCGAATACAGCTCACTGATCTTCACTGCGCT 660  
 QY 221 LeuLys\*\*\*LeuLysValGlyLysTyArgLysArgIleAspAspIleLeuAsnLysPhe 240  
 DB 661 TTGAAGATCTCAAGGTTGAAAGTATGAAAGAGATTTGATGATCTTGAACAACTTC 720  
 QY 241 AspProValValGluArgValIleLysLysArgArg\*\*\*IleValArgArgArg\*\*\*Asn 260  
 DB 721 GACCTGTGCTGAAAGGTCATCAAGAGCCGCTGATGATGTCAGAGAGAAAGAAC 780  
 QY 261 GlyGlu\*\*\*\*GluGlyGlu\*\*\*SerGlyVal\*\*\*LeuAspThrLeuLeuGluPheAla 280  
 DB 781 GAGAGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 QY 281 GluAspGluThr\*\*\*GluIleLysIleThrLys\*\*\*\*IleLysGlyLeuValValAsp 300  
 DB 841 GAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900  
 QY 301 \*\*\*PheSerAlaGly\*\*\*AspSerThrAla\*\*\*\*ThrGluTrpAlaLeuAlaGluLeu 320  
 DB 901 TTTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960  
 QY 321 IleAsnAsnPro\*\*\*ValLeu\*\*\*\*AlaArgGluGlu\*\*\*TySerValValGlyLys 340  
 DB 961 ATCAACATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020  
 QY 341 Asp\*\*\*LeuValAspGluValAspThrGlnAsnLeuProTyIleArgAlaIleValLys 360  
 DB 1021 GATGAGCTGTTGACGAGAGTGCACATCAAACTTCTTACATTTAGGCGCATTTGCAAG 1080  
 QY 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys\*\*\*GluGluCys 380  
 DB 1081 GAGACATTCGAGATGACCCACCACTCCAGTGGTCAAAAGAAAGTGCAGAGAGAGTGT 1140

QY 381 \*\*\*IleAsnGly\*\*\*Val\*\*\*ProGluGlyAlaLeu\*\*\*\*\*PheAsnValTrpGluVal 400  
 DB 1141 GAGATTAAATGAGATATGATATCCAGAGGAGATTTGTTCTTTCAATGTTGGCAAGTA 1200  
 QY 401 Gly\*\*\*Asp\*\*\*LysTyTrpAspArgProSerGlu\*\*\*ArgProGluArgPheLeuGlu 420  
 DB 1201 GGAAGAGCCCAATATCTGGGAGACAGACATCATGATTCCTCCGAGAGGTTCTTGA 1260  
 QY 421 Thr\*\*\*AlaGluGlyGluAla\*\*\*\*LeuAspLeuArgGly\*\*\*HisPheGlnLeuLeu 440  
 DB 1261 ACTGGCGCTGAAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 QY 441 ProPheGlySerGlyArg\*\*\*MetCysProGlyVal\*\*\*LeuAlaThrSerGly\*\*\*\*Ala 460  
 DB 1321 CCATTGGGCTTGGGAGAGATGTGCGCTGCTGCAATTTGGCTTACTTCAAGAAATGGA 1380  
 QY 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGlyGln 480  
 DB 1381 ACATCTTTCATCTTCTTATCAATGCTTTGACCTGCAAGTGTGAGGCTTCAAGACAA 1440  
 QY 481 IleLeuLysGly\*\*\*AspAlaLysValSerMetGluGluArgAlaGlyLeuThrValPro 500  
 DB 1441 ATTTGAAAGGTACAGATCCAAAGTTAGCATGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
 QY 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520  
 DB 1501 AGGCACTATGTCTCGTTTGTCTTCCACTTGAAGAGATCGGCGGTGCATTAACCTCT 1560  
 QY 521 Ser 521  
 DB 1561 TCT 1563

RESULT 2  
 AF195798  
 LOCUS AF195798 1774 bp mRNA linear PLN 16-FEB-2000  
 DEFINITION Glycine max isoflavone synthase 1 (Ifs1) mRNA, complete cds.  
 ACCESSION AF195798  
 VERSION AF195798.1 GI:6979519  
 KEYWORDS  
 SOURCE Glycine max.  
 ORGANISM Glycine max.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 REFERENCE 1 (bases 1 to 1774)  
 AUTHORS Jung W., Yu O., Lau S.M., O'Keefe D.P., Odell J., Fader G. and McConigle B.  
 TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes  
 JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)  
 MEDLINE 20124255  
 PUBMED 10657130  
 REFERENCE 2 (bases 1 to 1774)  
 AUTHORS Jung W., Yu O., Odell J., Fader G. and McConigle B.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-OCT-1999) Nutrition and Health, DuPont, P. O. Box





Db 301 TTCACACAGGTTCCAAACCTCAGCATMAGACGCTCAGCTATGATAGCTGAGTGGCC 360

Qy 121 \*\*\*\*\*Pro\*\*\*GlyProTyrTrp\*\*\*PheValArgLysLeuIleMetAsnAspLeu 140

Db 361 ATGTTCTCTCGACCTTACTGAGAGTGTGGAGAGGCTCATGATCAAGACCTTTC 420

Qy 141 AsnAlaThrThrValAsn\*\*\*LeuArgProLeuArgThrGlnGlnIleArgLys\*\*\*Leu 160

Db 421 AACGGCACACGCTTAAACAGATGAGGCTTTGAGACCCACACAGATCCGACAGTTCTT 480

Qy 161 Arg\*\*\*MetAlaGln\*\*\*AlaGluAla\*\*\*LysProLeuAsp\*\*\*ThrGluLysLeu 180

Db 481 AGGTTATGGCCCAAGGGCGAGAGGACAGAACCCCTTACCTTACCCAGAGACTTCG 540

Qy 181 LysTrp\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGlyLysLysLysLysLysLys 200

Db 541 AATATGACCAACAGACCATCTCATGATGATGCTCGGCGAGGCTGAGAGATCAAGAGAC 600

Qy 201 IleAlaArgLysValLeuLysIle\*\*\*GlyLysTrpSerLeuThrAspPheIle\*\*\*Pro 220

Db 601 ATGCTGCGCAGGCTTCTTAAAGATCTTTGGCGAATACAGCTCAGCTTCACTGAGCCA 660

Qy 221 LeuLys\*\*\*LeuLysValGlyLysTrpGlyLysArgIleAspAspIleLeuAsnLysPhe 240

Db 661 TTGAAGCATCTCAAGGTTGAGAGATGAGAGAGAGATCGACATCTTGAACAAAGTTC 720

Qy 241 AspProValValGluArgValIleLysLysArgArg\*\*\*IleValArgArg\*\*\*Asn 260

Db 721 GACCTGCTGCTTGAAGAGATCATCAAGACCCCTGAGATCTGAGAGAGAGAGAGAAC 780

Qy 261 GlyLys\*\*\*\*\*GluGlyGlu\*\*\*SerGlyVal\*\*\*LeuAspThrLeuGluPheAla 280

Db 781 GAGAGGTTGTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGGTTGAGG 840

Qy 281 GluAspGluThr\*\*\*GluIleLysIleThrLys\*\*\*IleLysGlyLeuValAsp 300

Db 841 GAGATAGACCATGAGATCAAAATCAACAGACCATCAAGGCTTGTGTCGAC 900

Qy 301 \*\*\*PheSerAlaGly\*\*\*AspSerThrAla\*\*\*\*\*ThrGluTrpAlaLeuAlaGluLeu 320

Db 901 TTTTCTCGACAGAACAGACTCCACAGCGGTGGACACAGAGTGGCATTTGGCAACTC 960

Qy 321 IleAsnAsnPro\*\*\*ValLeu\*\*\*\*\*AlaArgGluGlu\*\*\*TyrSerValValGlyLys 340

Db 961 ATCAACATCTTAAAGGTTGTGAAAGGCTGTGAGAGGCTTACAGTGTGTGGAAAG 1020

Qy 341 Asp\*\*\*LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360

Db 1021 GACAGACTGTGAGAGAGATGACACTCAAAACCTTCTTACATTAAGCAATCGAAG 1080

Qy 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys\*\*\*GluLysCys 380

Db 1081 GAGCATTTCCGACATGACCCCGACCTCCAGGTGCTCAAAAGAGATCGACAGAGAGTGT 1140

Qy 381 \*\*\*IleAsnGly\*\*\*Val\*\*\*ProGluGlyLysLeu\*\*\*\*\*PheAsnValTrpGlnVal 400

Db 1141 GAGATTAATGATATGATGATCCAGAGGAGCATTTGCTCTTCAATATGAGCAAGA 1200

Qy 401 Gly\*\*\*Asp\*\*\*LysTrpTrpAspArgProSerGlu\*\*\*ArgProLysArgPheLeuGlu 420

Db 1201 GGAAGAGACCCCAATATCTGAGACAGACATCGAGTTCTGAGAGGTTCTTAAAG 1260

Qy 421 Thr\*\*\*AlaGluGlyGluAla\*\*\*\*\*LeuAspLeuArgGly\*\*\*HisPheGlnLeuLeu 440

Db 1261 ACGGGGCTTAAGGGAGAGAGAGGCTTGTATCTTAAAGGAGACACATTTCAACTTTC 1320

Qy 441 ProPheGlySerGlyArg\*\*\*MetCysProGlyVal\*\*\*LeuAlaThrSerGly\*\*\*Ala 460

Db 1321 CCATTGGGCTGTGAGAGAGAGATGTGCCCTGAGATCATCTGACTTACCTCGGAATGGCA 1380

Qy 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGlyGln 480

Db 1381 ACACTTCTCGCATCTTATTCAGTGTCTTGAAGTGTGAGTGTGCTGACAGAGCAG 1440

Qy 481 IleLeuLysGly\*\*\*AspAlaLysValSerMetGluLysArgLysLysLeuThrValPro 500

Db 1441 ATATTGAAGGCTGTGATCGCCCAAGTTAGCATGAGAGAGAGAGAGAGAGAGAGAGAG 1500

Qy 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520

Db 1501 AGGGCATATGCTCTTGTGTGTTCACCTTGCAGAGATGGCGCTGACATTAAGTCTT 1560

Qy 521 Ser 521

Db 1561 TCT 1563

RESULT 4

AF195807 1566 bp mRNA linear PLN 16-FEB-2000

LOCUS AF195807

DEFINITION Vigna radiata isoflavone synthase 2 (lfst2) mRNA, complete cds.

ACCESSION AF195807

VERSION AF195807.1 GI:6979537

KEYWORDS

SOURCE Vigna radiata.

ORGANISM Vigna radiata.

REFERENCE 1 (bases 1 to 1566)

AUTHORS Jung, W., Yu, O., Lau, S. M., O'Keefe, D. P., Odell, J., Fader, G. and McGonigle, B.

TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes

JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)

MEDLINE 20124255

PUBMED 10657130

REFERENCE 2 (bases 1 to 1566)

AUTHORS Jung, W., Yu, O., Odell, J., Fader, G. and McGonigle, B.

TITLE Direct Submision

JOURNAL Submitted (18-OCT-1999) Nutrition and Health, Dupont, P. O. Box 80402, Wilmington, DE 19880-0402, USA

FEATURES

location/Qualifiers

source 1..1566

/organism="Vigna radiata"  
/db\_xref="taxon:157791"  
1..1566  
/gene="ifs2"  
1..1566  
/note="cytochrome P450"  
/product="isoflavone synthase 2"  
/protein\_id="AAF34528.1"  
/db\_xref="GI:6979538"  
/translation="MILELALGLVLAFLRLPPTAKSKALRLPSPSPKRLPF  
IGHILKDLHLVALIDLSKKGPLSLYSGSPVLYASSTBELFKFLQTHKATFN  
TRPTSAIRRLTYDSSVAMVPPGPKYKFRKLIMDLAATTNKLRLTQITKFL  
RAMAQGAQKPLDITELKWNSTISMMGLGAEIRDLAEVLKIFGYSITDI  
NPKLKLKVKYKRIIDILNKDFVVERVYKRIEIVRRKGEVGEVSGLDTL  
LEFADETIKITKHDIKGLVDFEAGTISVIAEMALDELINPKVLEKAREEV  
YSVQKRLVDEVDTONLPYRAIVKTEFRMPLPVYKRCCTECCEINGVYI  
ILFVWQVGRDPKYMDRSEPRPERLETGAGGARPDLRQHFQILPGRSRMCP  
GNLATSGMTLILSLIQCPLDVLGPQGLIKGDKAKVMEERAGLTVPRASLVCV  
PLARIGVASKILS"

BASE COUNT 414 a 393 c 387 g 372 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.03e-259 Length: 1566  
Score: 2268.00 Matches: 454  
Percent Similarity: 87.14% Conservative: 0  
Best Local Similarity: 87.14% Mismatches: 67  
Query Match: 94.66% Indels: 0  
D: 8 Gaps: 0

US-09-857-581-66 (1-521) x AF195807 (1-1566)

OY 1 MetIeuLeuGIuLeuAlaLeuGIyLeu\*\*\*ValIeuAlaLeuPhe\*\*\*HisIeuAlaArgPro 20  
DB 1 ATGTTGCTGGAACCTTGACCTTGTTATGCTTGTGCTCTGTTCTGCACTTGCGCTCC 60  
OY 21 ThrPro\*\*\*Ala\*\*\*SerIyAlaLeuArgHisIeuProAsnProProSerPro\*\*\*Pro 40  
DB 61 ACACCCACTGCMAATCAAAAGCATGCCATCTCCCAACCCACCAAGCCCAAGCCCT 120  
OY 41 ArgLeuProPheIleGIyHis\*\*\*HisIeuLeuIyAspIyLeuHisIyAla\*\*\* 60  
DB 121 CGTCTCCCTTCAITAGACACCTTCACTTAAAGACAACTTCTCCACGCGCTC 180  
OY 61 IleAspLeuSerIyHisIyGIyProLeuPheSer\*\*\*\*\*PheGIySerMetProThr 80  
DB 181 ATCGACCTCTCCAAAAAACAATGCTCCCTTATCTCTCTACTTGGCTCATGCAACC 240  
OY 81 ValIAlaIAserThrProGIuLeuPheIyLeuPheLeuGIu\*\*\*\*\*GIuAlaThrSer 100  
DB 241 GTTGTTCCTCCACACCAAGATGTTCAAGCTTCTCCCAACCAACGAGGCACTTCC 300  
OY 101 Phe\*\*\*ThrArgPheGIuThrSerAla\*\*\*Arg\*\*\*LeuThrIyAsp\*\*\*\*\*ValAla 120  
DB 301 TTCAACACAGAGTTCCAACTCAAGCCATAGACCTCACTATGATAGCTAGTGGCC 360

OY 121 \*\*\*\*\*Pro\*\*\*GIyProIyTrp\*\*\*PheValArgIyLeuIleMetAsnAspLeuLeu 140  
DB 361 ATGTTCCCTTCGACCTTACCTTGAACATGTTGAGAACGCTCACTAGAACACTTCTCC 420  
OY 141 AsnAlaThrThrValAsn\*\*\*LeuArgProLeuArgThrGIuInIleArgIyS\*\*\*Leu 160  
DB 421 AACGCCACCACTTAAACAGTGAAGGCTTGAAGACCCACCAACAGATCCGAACTTCTC 480  
OY 161 Arg\*\*\*MetAlaGIu\*\*\*AlaGIuAla\*\*\*LysProLeuAsp\*\*\*ThrGIuGIuLeuLeu 180  
DB 481 AGGCGTATGCGCCACAGGCGACAGGACCAAGAGCCCTTACCTTACCGAAGACTTCTG 540  
OY 181 LysTrp\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGIyAlaGIuGIuIleArgAsp 200  
DB 541 AATGACCAACACAGACATCTTCATATGATGCTCGGACAGGCTGAGGACATCAAGAC 600  
OY 201 IleAlaArgGIuValIeuIySalle\*\*\*GIyGIuIySerIeuThrAspPheIle\*\*\*Pro 220  
DB 601 ATCGCTCCGACAGGTTCTTAAAGATCTTGGGGAATCAAGCTTCACTGACTTCACTGCGCA 660  
OY 221 LeuIyS\*\*\*LeuIyValGIyIySArgIyLysArgIleAspAspIleLeuAsnIyS Phe 240  
DB 661 TTGAACATCTCAAGGTTGGAAGATGAGAGAGATCAAGACATCTTGAACAGTTC 720  
OY 241 AspProValIAlaGIuArgValIleIySArgArg\*\*\*IleValArgArgArg\*\*\*Asn 260  
DB 721 GACCTGTCGTGTAAGAGTCAATCAAGAGCCCTGATGCTGAGAGAGAGAGAGAAC 780  
OY 261 GIyGIu\*\*\*\*\*GIuGIyGIu\*\*\*SerGIyVal\*\*\*LeuAspThrLeuLeuGIuPheAla 280  
DB 781 GAGAGGTTGTTGAGGTGAGGTGAGGTGAGGAGGTTTCTTGACATCTTGCTGAATTCGCT 840  
OY 281 GIuAspGIuThr\*\*\*GIuIleIySleThrIyS\*\*\*\*\*IleIyGIyLeuValIAlaAsp 300  
DB 841 GAGATGAGACATGAGATGAGATCAAAATCAACAGACACATCAAGGCTTGTGCGAC 900  
OY 301 \*\*\*PheSerAlaGIy\*\*\*AspSerThrAla\*\*\*\*\*ThrGIuThrAlaLeuAlaGIuLeu 320  
DB 901 TTTTCTCGGACAGAACATCACTCCACGCGTGGCAGACAGATGGGCAATGGCAATCTC 960  
OY 321 IleAsnAsnPro\*\*\*ValIeu\*\*\*\*\*AlaArgGIuGIu\*\*\*TyrSerValIAlaGIyLys 340  
DB 961 ATCAACATCTTAAGTGTGTAAGAGGCTCGTGAAGAGGTCTACAGTGTGGGAAAG 1020  
OY 341 Asp\*\*\*LeuValIAspGIuValAspThrGIuAsnLeuProIyIleArgAlaIleValIyS 360  
DB 1021 GACAGACTTGGACCAAGTTCGACTCAAAACCTTCTTACATAGAGCAATCGTGAAG 1080  
OY 361 GIuThrPheArgMetHisProProLeuProValIAlaIySArgIyCys\*\*\*GIuGIuCys 380  
DB 1081 GAGACATTCGCAATGACCCGCACTCCGAGTGTCAAAAGAAAGTGAACGAGAGGTGT 1140  
OY 381 \*\*\*IleAsnGIy\*\*\*Val\*\*\*ProGIuGIyAlaLeu\*\*\*\*\*PheAsnValIArgIuVal 400  
DB 1141 GAGATTATGATATATATATATCCAGAGGAGCATTTGATCTTCAATGTATAGCAAGTA 1200

QY 401 Gly\*\*\*Asp\*\*\*Lys\*\*\*TrpAspArgProSerGlu\*\*\*ArgProGluValPheLeuGlu 420  
DB 1201 GGAAAGAGCCCAAAATACGGGACACACATGGAGATTCCTGCTGAGAGCTTACAG 1260  
QY 421 Thr\*\*\*AlaGluGlyGluAla\*\*\*\*\*LeuAspLeuArgGly\*\*\*HisPheGluLeuLeu 440  
DB 1261 ACAGGGGCTGAGAGGAGAGAGAGGCTTATCTTATGGGAGACAACTTTCAACTTCTC 1320  
QY 441 ProPheGlySerGlyArg\*\*\*MetCysProGlyVal\*\*\*LeuAlaThrSerGly\*\*\*Ala 460  
DB 1321 CCATTGGGCTGCGAGAGAGATGCTCGAGATCATCTGCTACTTCCGAGATGGCA 1380  
QY 461 ThrLeuLeuAlaSerLeuLeuLeuGlyCysPheAspLeuGluValLeuGlyProGluGlyGlu 480  
DB 1381 AACTCTTCGATCTCTTATTCAGCTGCTTTCAGAGATGCTGGGCTCACAAGAGCAG 1440  
QY 481 IleLeuGlyGly\*\*\*AspAlaValSerMetGluGluArgAlaGlyLeuThrValPro 500  
DB 1441 ATATTGAGGGTGTGTACCCCAAGATTAGCATGAGAGAGAGAGGAGGCTTCACTGTTCA 1500  
QY 501 ArgAlaHisSerLeuValCysValProLeuAlaArgGlyGlyValAlaSerGlyLeuLeu 520  
DB 1501 AGGAGCATAGCTGTGTCTGTCTTCCACTTTCAGAGATCCGCTTGCATCTAACTCTT 1560  
QY 521 Ser 521  
DB 1561 TCT 1563  
RESULT 5  
AF195808  
LOCUS AF195808 1566 bp mRNA linear PLN 16-FEB-2000  
DEFINITION Vigna radiata isoflavone synthase 3 (if3) mRNA, complete cds.  
ACCESSION AF195808  
VERSION AF195808.1 GI:6979539  
KEYWORDS  
SOURCE Vigna radiata.  
ORGANISM Vigna radiata  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Vigna.  
REFERENCE  
AUTHORS 1 (bases 1 to 1566)  
Jung, W., Yu, O., Lau, S.M., O'Keefe, D.P., Odell, J., Fader, G. and  
McGonigle, B.  
TITLE Identification and expression of isoflavone synthase, the key  
enzyme for biosynthesis of isoflavones in legumes  
JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)  
MEDLINE 20124255  
PUBMED 10657130  
REFERENCE  
AUTHORS 2 (bases 1 to 1566)  
Jung, W., Yu, O., Odell, J., Fader, G. and McGonigle, B.  
TITLE Direct Submission  
JOURNAL Submitted (18-OCT-1999) Nutrition and Health, DuPont, P. O. Box  
80402, Wilmington, DE 19880-0402, USA  
FEATURES  
source 1..1566  
location/Qualifiers  
/organism="Vigna radiata"

gene /db xref="taxon:157791"  
1..1566  
/gene="if3"  
1..1566  
/gene="if3"  
/note="cytochrome P450"  
/codon\_start=1  
/product="isoflavone synthase 3"  
/protein\_id="AF34529.1"  
/translation="MLELALGLVLAFLRLPPTPAKSKALRLHPSPKPRPLP  
IGHILKDKLHVALIDLSKKGPLESLYRSGMPTVAATPELPLQTHETSFN  
TRPQSAIRRLITDSSVAMPFGYKFKPKLIMNDLNAITVANKRLKQIRKFL  
RVMAQGAAPQPLDLTEILKMTNSTISMMWLGAESEIRIDIAEVLKIFGYSILDFI  
WPKLKVKKYKRIIDILNKPDPVVERVYKRRREIVRRKNGEVBEVSGVFLDL  
LEFADETTEIKITKDIKGLVDFPSAGTSDTAVATEMLAELINPKVLEKAREV  
YAVVKDLNDEVDTNQLPYIRALIVETRMHPLPVVKKRCCECEINGVYIPCAL  
ILFNWQVGRPKRMRSERPERLETGAGEARPLDRQHQLPLPSSGRNCP  
GNLATSMTLASLIQCFPLQVLGPQQLKGGDAKVSMEERAGLTVPASLVCV  
PLARIGVASKLS"  
BASE COUNT 415 a 393 c 386 g 372 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.03e-259 Length: 1566  
Score: 2268.00 Matches: 454  
Percent Similarity: 87.14% Conservative: 0  
Best Local Similarity: 87.14% Mismatches: 67  
Query Match: 94.66% Indels: 0  
DB: 8 Gaps: 0  
US-09-857-581-66 (1-521) x AF195808 (1-1566)  
QY 1 MetLeuLeuGluLeuAlaLeuGlyLeu\*\*\*ValLeuAlaLeuPhe\*\*\*HisLeuArgPro 20  
DB 1 ATGTTCGAGACTTCGACTTGGTTTATGTTGGTCTGTTCTGCACTTCGCTCC 60  
QY 21 ThrPro\*\*\*Ala\*\*\*SerValAlaLeuArgHisLeuProAsnProSerPro\*\*\*Pro 40  
DB 61 ACACCCACTGCAAAATCAAAAGCATTCGCAATCTCCAAACCAACCAAGCCCAAGCCCT 120  
QY 41 ArgLeuProPheIleGlyHis\*\*\*HisLeuLeuValAspIleuLeuHisTyrAla\*\*\* 60  
DB 121 CGCTTCCTTCATAGACACTTCATCTTAAAGACAAATCTTCACATGACATC 180  
QY 61 IleAspLeuSerIleGlyHisGlyProLeuPheSer\*\*\*\*\*PheGlySerMetProThr 80  
DB 181 ATGACACTCTCCAAAACATAGTCCCTTATCTCTCTACTTGGCTCAGTCCAGCC 240  
QY 81 ValValAlaSerThrProGluLeuPheLeuLeuLeuGlu\*\*\*\*\*GluAlaThrSer 100  
DB 241 GTTGTTCCTTCACACAGAAATGTCAGCTTCCTCCAAACGACAGAGCAACTTCC 300  
QY 101 Phe\*\*\*ThrArgPheGluThrSerAla\*\*\*Arg\*\*\*LeuThrTyrAsp\*\*\*\*\*ValAla 120  
DB 301 TTCAACACAGTTCCAAACTTCAGCCATAGAGCGCTCACTATGATGCTCAGTGGCC 360



QY 121 \*\*\*\*\*Pro\*\*\*GlyProTyrTrp\*\*\*PheValArgLysLeuIleMetAsnAspLeuLeu 140  
 DB 361 ATGCTTCCTTGGACCTTACTGAAAGTTCTGAGAGAGCTCATATAGACGCTTCTC 420  
 QY 141 AsnAlaThrValAsn\*\*\*LeuArgProLeuArgThrGlnIleArgLys\*\*\*Leu 160  
 DB 421 AACGACACCTGTAAACAGATTGAGCCTTTGAGACCCAAACATATCCGAGTTCTT 480  
 QY 161 Arg\*\*\*MetAlaGln\*\*\*AlaGluAla\*\*\*LysProLeuAsp\*\*\*ThrGluLeuLeu 180  
 DB 481 AGGTTATGGCCCAAGCGCAGAGGACAGAGCCCTTGTACTGACGAGAGCTTCTG 540  
 QY 181 LysTrp\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGlyGluAlaGluIleLeuAsp 200  
 DB 541 AAATGACCAACAGCATCTCATGATGATGCTCGCGAGCTGAGAGATCAAGAC 600  
 QY 201 IleAlaArgIleValLeuLysIle\*\*\*GlyGluTyrSerLeuThrAspPheIle\*\*\*Pro 220  
 DB 601 ATCGCTGCGAGGTTCTTATGATCTTGGCGAATACAGCTCTCATCTCATCTGCGCA 660  
 QY 221 LeuLys\*\*\*LeuLysValGlyLysTyrGluLysArgIleAspAspIleLeuAsnLysPhe 240  
 DB 661 TTGAAGCATCTCAAGTTGGAAGTATGAGAGAGATGACAGATCTTGAACAAAGTTT 720  
 QY 241 AspProValValGluArgValIleLysLysArgArg\*\*\*IleValArgArgArg\*\*\*Asn 260  
 DB 721 GACCTGCTGTTGAAAGATCATCAAGAGCGCGTGAATGCTGAGAGAGAGAAAGAC 780  
 QY 261 GlyGluLys\*\*\*GlyGlyGluLys\*\*\*SerGlyVal\*\*\*LeuAspThrLeuGluLysPhe 280  
 DB 781 GAGAGGTTGTTGAGGAGGTGACGCGGCGTTCCTTGACATCTTCTTGAAATCGCT 840  
 QY 281 GluAspGluThr\*\*\*GluIleLysIleThrLys\*\*\*IleLysGlyLeuValIleAsp 300  
 DB 841 GAGATGACACACGAGATCAAAATACCAAGACACATCAAGAGGCTTGTGTGAC 900  
 QY 301 \*\*\*PheSerAlaGly\*\*\*AspSerThrAla\*\*\*ThrGluTyrAlaLeuAlaGluLeu 320  
 DB 901 TTTTCTCGGAGAAACAGACTCAAGCGGTGGCAACAGATGGGCAATTGGCAAGATC 960  
 QY 321 IleAsnAsnPro\*\*\*ValLeu\*\*\*AlaArgGluGlu\*\*\*TyrSerValIleGlyLys 340  
 DB 961 ATCAACATCTTAAAGTGTGTGAAAGGCTGTGAGAGAGGCTTACAGGTTGGGAAAG 1020  
 QY 341 Asp\*\*\*LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360  
 DB 1021 GACGAGCTTGTGAGAGAGTTGACCTCAAAACCTTCTTACATTAGAGCAATCGGAAG 1080  
 QY 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys\*\*\*GluGluCys 380  
 DB 1081 GAGCATTTCCGATGACCCGCGCATCTCCAGGTGTCAAAAGAAAGTGTGACAGAGAGTGT 1140  
 QY 381 \*\*\*IleAsnGly\*\*\*Val\*\*\*ProGluGlyAlaLeu\*\*\*PheAsnValTyrGlnVal 400  
 DB 1141 GAGATTAATGAGATGATGATCCAGAGAGCATTTGATTTCTTCAATGATGCGAAGTA 1200  
 QY 401 Gly\*\*\*Asp\*\*\*LysTyrTrpAspArgProSerGlu\*\*\*ArgProGluArgPheLeuGlu 420

DB 1201 GGAAGAGACCCCAATATCTGGGACAGACATCGAGTTCCGTCCTTGAAGGTTCTTGAG 1260  
 QY 421 Thr\*\*\*AlaGluGlyGluAla\*\*\*LeuAspLeuArgGly\*\*\*HisPheGlnLeuLeu 440  
 DB 1261 ACAGGGGCTGAAAGGAG 1320  
 QY 441 ProPheGlySerGlyArg\*\*\*MetCysProGlyVal\*\*\*LeuAlaThrSerGly\*\*\*Ala 460  
 DB 1321 CCATTGGGTCTGCGAGAGAGATGTGCTCGAGATGATCTGCTACTTCCGGAATGGCA 1380  
 QY 461 ThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValLeuGlyProGlnGlyGln 480  
 DB 1381 ACATCTTGCATCTTATTAAGTGTGACTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTG 1440  
 QY 481 IleLeuLysGly\*\*\*AspAlaLysValSerMetGluGluArgAlaGlyLeuThrValPro 500  
 DB 1441 ATATTGAAAGGTGTGTGACCCCAAGTTAGCATGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
 QY 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520  
 DB 1501 AGGGCAATATGCTGTGTGTGTTCCACTTGCAGAGATCGGCGTTCATTAATCTCTT 1560  
 QY 521 Ser 521  
 DB 1561 TCT 1563

# RESULT 6

AF195810 1566 bp mRNA linear PLN 16-FEB-2000  
 LOCUS Trifolium pratense isoflavone synthase 1 (ifsi) mRNA, complete cds.  
 DEFINITION AF195810  
 ACCESSION AF195810  
 VERSION AF195810.1 GI:6979543

KEYWORDS Trifolium pratense.  
 SOURCE Trifolium pratense.  
 ORGANISM Trifolium pratense.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Trifolium.

REFERENCE 1 (bases 1 to 1566)

Jung, W., Yu, O., Lau, S.M., O'Keefe, D.P., Odell, J., Fader, G. and McConigle, B.  
 Identification and expression of isoflavone synthase, the key

enzyme for biosynthesis of isoflavones in legumes  
 Nat. Biotechnol. 18 (2), 208-212 (2000)

JOURNAL MEDLINE 20124255

PUBMED 10657130

REFERENCE 2 (bases 1 to 1566)  
 Fader, G. and McConigle, B.  
 Direct Submision  
 Submitted (18-OCT-1999) Nutrition and Health, Dupont, P.O. Box  
 80402, Wilmington, DE 19880-0402, USA

FEATURES Location/Qualifiers

1..1566  
 /organism="Trifolium pratense"  
 /db\_xref="taxon:57577"

gene 1..1566  
/gene="1fsl"  
1..1566  
CDS /note="cytochrome P450"  
/codon\_start=1  
/product="isoflavone synthase 1"  
/protein\_id="AA134531.1"  
/db\_xref="GI:6979544"

BASE COUNT 416 a 393 c 387 g 370 t  
ORIGIN  
1 MetLeuLeuGluLeuAlaLeuGlyLeu\*\*\*ValLeuAlaLeuPhe\*\*\*HisLeuArgPro 20  
21 ThrPro\*\*\*Ala\*\*\*SerLySAlaLeuArgHisLeuProAsnProProSerPro\*\*\*Pro 40  
61 ACACCCACTGCAGAAATCAAAAGACCTTCGCATCCCAACCCCAACCAAGCCCAAGCCT 120  
41 ArgLeuProPheIleGlyHis\*\*\*HisLeuLeuLysAspLysLeuLeuHisTyrAla\*\*\* 60  
121 CGCTCTCCCTTCATAGACACCTTCATCTCTTAAAGACAACTCTCCACCTACGCACTC 180  
61 IleAspLeuSerLysHisGlyProLeuPheSer\*\*\*\*PheGlySerMetProThr 80  
181 ATGACCTCTCCCAAAACATGATGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln\*\*\*\*GluAlaThrSer 100  
241 GTTGTGCTCCACACACAGAAATTTCTCAAGCTCTTCTCCAAACGACACGACGCACTTCC 300  
101 Phe\*\*\*ThrArgPheGlnThrSerAla\*\*\*Arg\*\*\*LeuThrTyrAsp\*\*\*\*ValAla 120  
301 TTCACACAGAGTTCACAACTCAGCCATAGACGCTCAGCTATAGCTAGCTGAGCC 360  
121 \*\*\*\*\*Pro\*\*\*GlyProTyrTrrp\*\*\*PheValArgLysLeuIleMetAsnAspLeu 140

Alignment Scores:  
Pred. No.: 1,03e-259 Length: 1566  
Score: 2268.00 Matches: 454  
Percent Similarity: 87.14% Conservative: 0  
Best Local Similarity: 87.14% Mismatches: 67  
Query Match: 94.66% Indels: 0  
DB: 8 Gaps: 0  
US-09-857-581-66 (1-521) x AF195810 (1-1566)

361 ATGCTCCCACTCGACACTTACGAAAGTTGTGAGAAAGCTCATCTAGACACTTCTC 420  
141 AsnAlaThrThrValAsn\*\*\*LeuArgProLeuArgThrGlnGlnIleArgLys\*\*\*Leu 160  
421 AACGCCACCACTGTAACAAAGTTGAGGCTTTGAGAGCCCAACAGATCCGCAACTTCTT 480  
161 Arg\*\*\*MetAlaGln\*\*\*AlaGluAla\*\*\*LysProLeuAsp\*\*\*ThrGluLeuLeu 180  
481 AGGCTTAATGAGCCCAAGCGCAGAGCCAGAGAGCCCTTCACTTGAACGAGACTTCTCG 540  
181 LysTrrp\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGlyGluAlaGlnGlnIleArgAsp 200  
541 AAATGACCAACAGCAACCACTTCATGATGATGATGCTGAGGCTGAGAGATGAGAGAC 600  
201 IleAlaArgGluValLeuLysIle\*\*\*GlyGluTyrSerLeuThrAspPheIle\*\*\*Pro 220  
601 ATGCTCCCGAGGTTCTTAAGTCTTTGCGGAAATACAGCTCCTCACTTCACTTCACTTCC 660  
221 LeuLys\*\*\*LeuLysValGlyLysTyrGluLysArgGlyLeuAspAspIleLeuAsnLysPhe 240  
661 TTGAGCATCTCAAGGTTGAAAGATGAGAGAGAGATCGACATCTTGAACAGATTTC 720  
241 AspProValValGluArgValIleLysLysArgArg\*\*\*IleValArgArgArg\*\*\*Asn 260  
721 GACCTCTCGTTGTAAGAGTATCAAGAGCCCGCTGATCTGAGAGAGAGAGAGAGAGAG 780  
261 GlyGluLys\*\*\*GluGlyGlu\*\*\*SerGlyVal\*\*\*LeuSerPheLeuLeuGluPheAla 280  
781 GAGAGAGTTGATAGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAGGT 840  
281 GluAspGluThr\*\*\*GluIleLysIleThrLys\*\*\*\*IleLysGlyLeuValValAsp 300  
841 GAGATGAGACCAAGAGATCAAAATCAACAGACACATCAAGGCTCTTGTTCGAC 900  
301 \*\*\*PheSerAlaGly\*\*\*AspSerThrAla\*\*\*\*ThrGluThrAlaLeuAlaGluLeu 320  
901 TTTTCTCGAGAGAGACAGATCCACAGCGGTGGACAGAGGTGGCATTTGGAGAACTC 960  
321 IleAsnAsnPro\*\*\*ValLeu\*\*\*\*AlaArgGluGlu\*\*\*TyrSerValValGlyLys 340  
961 ATCAACAAATCTTAAGCTGTGAAAGAGGCTCGTGAAGAGGTCTACAGTGTGTGGAAAG 1020  
341 Asp\*\*\*LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360  
1021 GACAGACTTGGAGACAGAGTGAACCTCAACCTTCTTCACTTGAAGCATGTGAAG 1080  
361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys\*\*\*GluGluCys 380  
1081 GAGACATTCGCGCATGACCCCGCCTCCGCTGCTCAAAAAGAAAGTGCACAGAGAGCTT 1140  
381 \*\*\*IleAsnGly\*\*\*Val\*\*\*ProGluGlyAlaLeu\*\*\*\*PheAsnValTrrpGlnVal 400  
1141 GAGATTAATGATATATGATCCAGAGAGGAGCATGATCTCTTCAAGTATGAGCAAGTA 1200  
401 Gly\*\*\*Asp\*\*\*LysTyrTrrpAspArgProSerGlu\*\*\*ArgProGluArgPheLeuGlu 420



Db 161 ATGGTTCCTTCGGAAGCTTACTGGAAGTTCTGAGAGAAAGCTTCATGAAAGCACTTCTC 420

Qy 141 AenAlaThrThyValAsn\*\*\*LeuAtrProLeuAtrThiGlnHlaGlyS\*\*\*Leu 160

Db 421 AACGCCACCACTGTAAACAGTTGAGGCTTTGAGAGCCCAAGATCCGAGATTCCTT 480

Qy 161 Arg\*\*\*MetAlaGlu\*\*\*AlaGluAla\*\*\*LysProLeuAsp\*\*\*ThrGluGluLeu 180

Db 481 AGGTTTATGACCCAGCGCAGAGGCAAGAAAGCCCTTACTTGAACGAGAGTTCTTG 540

Qy 181 LysTrp\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGlyGluAlaGluLysLeuAsp 200

Db 541 AAATGACCAACAGCAACATCTCATGATGATGCTCCGCGAGGCTGAGAGATCAAGAC 600

Qy 201 ILeaIaArgIuValLeuLysIle\*\*\*GlyGlyLysSerLeuThrAspPheIle\*\*\*Pro 220

Db 601 ATGCTGCGAGGTTCTTAATGATCTTGGCAATACAGCTCACTGATCTGATCTGCA 660

Qy 221 LeuLys\*\*\*LeuLysValGlyLysIleGlyLysArgIleAspAspIleLeuAsnLysPhe 240

Db 661 TTGAAGCATCTCAAGGTTGAAAGTATGAAAGAGATCCAGCATCTTGAACAAGTTTC 720

Qy 241 AspProValValGluArgValIleLysLysArgArg\*\*\*IleValArgArgArg\*\*\*Asn 260

Db 721 GACCTGTCTGTAAGAGATCATCAAGAGCCGCTGATGCTGAGAGAGAAAGAAC 780

Qy 261 GlyLysLys\*\*\*GluGlyGlu\*\*\*SerGlyVal\*\*\*LeuAspThrLeuLeuGluPheAla 280

Db 781 GAGAGGTTGTGAGAGGTGAGTCAAGGAGGTTTCTTACATCTTCTGAAATTCGCT 840

Qy 281 GluAspGluThr\*\*\*GluIleLysIleLysIleLysIleLysIleLysIleLysIleLys 300

Db 841 GAGATGAGACACGAGATCAAAATCACCAAGACATCAAGAGGTTGTTGTCGAC 900

Qy 301 \*\*\*PheSerAlaGly\*\*\*AspSerThrAla\*\*\*ThrGluTrpAlaLeuAlaGluLeu 320

Db 901 TTTTCTTCGACAGAAACACTCTCACAGCGTGGCAACAGATGGGCAATGGCAACTC 960

Qy 321 IleAsnAsnPro\*\*\*ValLeu\*\*\*AlaArgGluGlu\*\*\*TyrSerValValGlyLys 340

Db 961 ATCAACATCTTAAGGTGTGAAAGGCTCTGTAGAGAGGTCTTACAGTGTGTGGAAAG 1020

Qy 341 Asp\*\*\*LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360

Db 1021 GACAGACTGTGAGCAAGATTGACACTCAAAACCTTCTTACATTAGCAATGTGAAAG 1080

Qy 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys\*\*\*GluGluCys 380

Db 1081 GAGACATTCGATGACCGCCGCACTCCAGGTGCAAAAGAAAGTGCAGAAAGGTGT 1140

Qy 381 \*\*\*IleAsnGly\*\*\*Val\*\*\*ProGluGlyAlaLeu\*\*\*PheAsnValTrpGlnVal 400

Db 1141 GAGATTATGATGATGATCCAGAGGAGCATTTCTTCAATGATGCAAGTA 1200

Qy 401 Gly\*\*\*Asp\*\*\*LysTrpTrpAspArgProSerGlu\*\*\*ArgProGluArgPheLeuGlu 420

Db 1201 GGAAGAGACCCCAATATCGGACAGACATCGAGTTCTGCTCGAGAGGTCTCTAAG 1260

Qy 421 Thr\*\*\*AlaGluGlyGluAla\*\*\*LeuAspLeuArgGly\*\*\*HisPheGlnLeuLeu 440

Db 1261 AACGGGCTGAAGGGGAGAACAGGCTCTTGATCTTACGGGAGCAACATTTTCAACTTCTC 1320

Qy 441 ProPheGlySerGlyArg\*\*\*MetCysPheGlyVal\*\*\*LeuAlaThrSerGly\*\*\*Ala 460

Db 1321 CCATTGGGCTCTGGAGAGAGATGTGCTTGAGAGCATCTGACTTCTCGGAAATGCGCA 1380

Qy 461 ThrLeuAlaSerLeuIleGlnCysPheAspLeuGlnValIleuGlyProGlnGlyGln 480

Db 1381 AACCTTCTGATCTTATTCATGATGCTTTGACTTGAAGTGTGGTGCACAGACAG 1440

Qy 481 IleLeuLysGly\*\*\*AspAlaLysValSerMetGluLysArgAlaGlyLeuThrValPro 500

Db 1441 ATATTGAAGGCTGTACCCCAAGTTAGCATGGAAGAGAGGCGGCTCATCTGTTCA 1500

Qy 501 ArgAlaHisSerLeuValCysValProLeuAlaArgIleGlyValAlaSerLysLeuLeu 520

Db 1501 AGGCAACATGATCTTGTGTGTTCACCTTGCAAGATCGGCGTGCATCTTAACTTCTT 1560

Qy 521 Ser 521

Db 1561 TCT 1563

RESULT 8

AF195809 1566 bp mRNA linear PLN 16-FEB-2000

LOCUS AF195809

DEFINITION Vigna radiata isoflavone synthase 4 (lf4s) mRNA, complete cds.

ACCESSION AF195809

VERSION AF195809.1 GI:6979541

KEYWORDS

SOURCE Vigna radiata.

ORGANISM Vigna radiata.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.

REFERENCE 1 (bases 1 to 1566)

AUTHORS Jung,W., Yu,O., Lau,S.M., O'Keefe,D.P., Odell,J., Fader,G. and McConigle,B.

TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes

JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)

MEDLINE 20124255

PUBMED 10657130

REFERENCE 2 (bases 1 to 1566)

AUTHORS Jung,W., Yu,O., Odell,J., Fader,G. and McConigle,B.

TITLE Direct Submission

JOURNAL Submitted (18-OCT-1999) Nutrition and Health, Dufont, P.O. Box 80402, Wilmington, DE 19880-0402, USA

FEATURES

source

1. 1566

/organism="Vigna radiata"

/db\_xref="taxon:157791"

1. 1566

/gene="lf4s"

CDS

1..1566  
/gene="ife4"  
/note="cytochrome P450"  
/codon\_start=1  
/product="isoflavone synthase 4"  
/protein\_id="AAF34530.1"  
/db\_xref="GI:6979542"  
/translation="MLBELAGLLVLAFLHRLPTPAASKALRLPNPSPKRLPLF  
ICHLHLKDLKLVLAIDLSKKHGPLSLYFGSMPTVAATPELFKLTQTHATSFN  
TRFOTSAIRLTYDSSVAMVPGPYWKPVKRLINMDLMAITVNLRLPRLTQOIRKFL  
RVMQAQAQRKPLDLEELKMTNSTISMMMLGEAEIRIDIAEVLAKIFGEYSLDPI  
WPLKHLVKGVEKRIIDILINKPDPVERVYIKRRRIYVRKRGVEVEGSGVLEDTL  
LEFADETMEIKITKDHIKGLVDFPSAGTSTAEATWALMELINPKVLEKAEV  
YSVGGKORLVDEVDQNLPIYRAIVKETEPMPLPVKKKCTECCINGVILEGAL  
ILFNWQVGDPIKYMDRSEPERPERLETGAEARPLDRQHFQLLPFGSGRMCP  
GVNLATSGMATTILASLIQCFDLQVLPQQLIKGDAKVMERAGLTVPRASIVCV  
PLARIGVASKLTS"

BASE COUNT 417 a 392 c 385 g 372 t  
ORIGIN

Alignment Scores:

Pred. No.: 1,356-259 Length: 1566  
Score: 2267.00 Matches: 454  
Percent Similarity: 87.14% Conservative: 0  
Best Local Similarity: 87.14% Mismatches: 67  
Query Match: 94.62% Indels: 0  
DB: 8 Gaps: 0

US-09-857-581-66 (1-521) x AF195809 (1-1566)

QY 1 MetleuLeuGluLeuAlaLeuGlyLeu\*\*\*ValLeuAlaLeuPhe\*\*HisLeuArgPro 20  
DB 1 ATGTTGGTGGAACTTGCACTGGTTATATGGTTTGGCTGCTGTTCTGCACCTGGCTCCC 60  
QY 21 ThrPro\*\*Ala\*\*\*SerLyseAlaLeuArgHisLeuProAsnProProSerPro\*\*\*Pro 40  
DB 61 ACACCCCACTGCMAAAATCAAAAGCACTCCCATCTCCCAACCAACCAACCAAGCAAGCAAGCT 120  
QY 41 ArgLeuProPheIleGlyHis\*\*\*HisLeuLeuLysAspLysLeuLeuHisTyAla\*\*\* 60  
DB 121 GGTCTTCCCTTCATAGACACCTTCATCTTAAAGACAACTTCTCAGTACAGCACTC 180  
QY 61 IleAspLeuSerLysLysGlyProLeuPheSer\*\*\*\*PheGlySerMetProThr 80  
DB 181 ATGACCTCTCCAAAACATGCTGCTTATTTCTCTCTACTTTGGCTCCATGCAAGC 240  
QY 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln\*\*\*\*GluAlaThrSer 100  
DB 241 GTTGTGCTCCACACAGAAATGTTCAAGCTTCTCCCAAAAGCAAGCAAGCAAGCAAGCTTC 300  
QY 101 Phe\*\*\*ThrArgPheGlnThrSerAla\*\*\*Arg\*\*LeuThrTyArg\*\*\*\*ValAla 120  
DB 301 TTCACACCAAGGTTCCAAACCTCAGCCATTAAGACGCTCAGTATGATGCTCAGTGCC 360  
QY 121 \*\*\*\*\*Pro\*\*GlyProTyTrp\*\*\*PheValArgLysLeuLeuMetAsnAspLeu 140  
DB 361 ATGTTCCCTTGAGACTTACTGGAAGTTCTGAGGAAGCTCATCATGAAGCAAGCTTCTC 420

QY 141 AsnAlaThrThrValAsn\*\*\*LeuArgProLeuArgHisGlnGlnIleArgLys\*\*\*Leu 160  
DB 421 AAGGACCAACATCTATAACAGTGAAGCCCTTGAAGCCCAACAGATCCGCAAGTCTTCTT 480  
QY 161 Arg\*\*\*MetIleGln\*\*\*AlaGluAla\*\*\*LysProLeuAsp\*\*\*ThrGluGluLeuLeu 180  
DB 481 AGGATTATGCCCCAAGCGCAAGGCAAGAGCCCTTCACTTGAACCAAGCAAGCTTCTG 540  
QY 181 LysTrp\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGlyGluAlaGluLysIleAsp 200  
DB 541 AATATGACCAACAGCAACCACTCCTCATATGATGCTCCGCGAGGCTGAGAGATCAAGAC 600  
QY 201 IleAlaArgGluValLeuLysIle\*\*\*GlyGluTySerLeuThrAspPheIle\*\*\*Pro 220  
DB 601 ATGCTCCGAGGTTCTTAAGTCTTGAAGTCTTGAAGTCAAGCTTCACTGACTTCAAGTCCCA 660  
QY 221 LeuLys\*\*LeuLysValGlyLysTyArgLysArgGluIleAspAspIleLeuAsnLysPhe 240  
DB 661 TTGAAGCATCTCAAGCGTTGCAAGATGAGAGAGATGAGACATGCAAGCAAGCTTCAAGAGTTC 720  
QY 241 AspProValValGluArgValIleLysLysArgArg\*\*\*IleValArgArgArg\*\*\*Asn 260  
DB 721 GACCTGCTGCTTAAGAGATCAAGATCAAGAGCCCTGTAAGTCTGAGAGAGAGAGAGAGAC 780  
QY 261 GlyGlu\*\*\*\*\*GluGlyGlu\*\*SerGlyVal\*\*\*LeuAspThrLeuLeuGluPheAla 280  
DB 781 GGAAGAGCTTGTGAGGATGAGTCAAGGAGGTTTCTTCTTCACTTGGCTGAATTCGCT 840  
QY 281 GluAspGluThr\*\*GluIleLysIleThrLys\*\*\*\*\*IleLysGlyLeuValValAsp 300  
DB 841 GAGATGAGACATGAGATCAAAATCAACAGACCAATCAAGAGGCTTGTGTCGAC 900  
QY 301 \*\*\*PheSerAlaGly\*\*AspSerThrAla\*\*\*\*\*ThrGluTrpAlaLeuAlaGluLeu 320  
DB 901 TTTTCTCGGAGAGAACAGACTCCACAGCGAGGCAAGAGTGGCAATGGCAAGCTC 960  
QY 321 IleAsnAsnPro\*\*ValLeu\*\*\*\*\*AlaArgGluGlu\*\*TySerValValGlyLys 340  
DB 961 ATCAACAAATCTTAAGGTGTTGAAAGAGCTCGTGAAGAGGCTTACAGTGTGAGGAAG 1020  
QY 341 Asp\*\*\*LeuValAspGluValAspThrGlnAsnLeuProTyIleArgAlaIleValLys 360  
DB 1021 GACAGACTTGTGACAGAGTGAAGTCACTCAAAACCTTCTTACATTAGACAAATGTAAG 1080  
QY 361 GluThrPheArgMetHisProProLeuProValValLysArgLysCys\*\*GluGluCys 380  
DB 1081 GAGACATTCGATGACACCGGCACTCCAGTGTCAAAAGAAAGTGCAGAAAGTGT 1140  
QY 381 \*\*\*IleAsnGly\*\*Val\*\*\*ProGluGlyAlaLeu\*\*\*\*\*PheAsnValTrpGlnVal 400  
DB 1141 GAGATTAATGATATGATGATCCAGAGGAGCATTTGATCTTCAATGATATGCAAGTA 1200  
QY 401 Gly\*\*\*Asp\*\*\*LysTyTrpAspArgProSerGlu\*\*\*ArgProGluArgPheLeuGlu 420  
DB 1201 GGAAGAGACCCCAAAATCTGGAAGACATGAGAGTTCGTTCTGAGAGAGTTCATAGAG 1260



Db 96 ACACCACTGCAMAAATCAAAAGCATCTCCATCTCCAAACCCACCAAGCCCAAGACT 155  
 Qy 41 ArgLeuProthe11eglyHis\*\*\*HisLeuLeuLysAspLysLeuLeuHis15Trala\*\*\* 60  
 Db 156 CCGTCTCCCTCTCAAGACACCTTCACTCTTAAAGACAACTCTCCATCGACGACCTC 215  
 Qy 61 IleAspLeuSerLysHisGlyProLeuPheSer\*\*\*\*\*PheGlySerMetProthr 80  
 Db 216 ATCGACCTCTCAAAAACATGATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 275  
 Qy 81 ValValAlaSerThrProGluLeuPheLysLeuPheLeuGln\*\*\*\*\*GluAlaThrSer 100  
 Db 276 GTTGTGCTCTCAACACCAATTTCTCAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 335  
 Qy 101 Phe\*\*\*ThrArgPheGlnThrSerAla\*\*\*Arg\*\*\*LeuThrTyrAsp\*\*\*\*\*ValAla 120  
 Db 336 TTCACACAGAGTTCACAACTCAAGCATTAAGACCTCTCACTATGATGCTGAGGCTC 395  
 Qy 121 \*\*\*\*\*Pro\*\*\*GlyProTyrTTP\*\*\*PheValArgLysLeuLemecAsnAspLeuLeu 140  
 Db 396 ATGCTTCCCTCGACCTTACTGAGAGTTCGAGAGGATCTCATCATCAAGACTTCTCC 455  
 Qy 141 AsnAlaThrThrValAsn\*\*\*LeuArgProLeuArgThrGlnGlnLysLys\*\*\*Leu 160  
 Db 456 AACGCACACCTGTAACAGAGTGAAGCTTGAAGACCCACACATCCGACAGTTCTT 515  
 Qy 161 Arg\*\*\*MetAlaGln\*\*\*AlaGluAla\*\*\*LysProLeuAsp\*\*\*ThrGluLysLeuLeu 180  
 Db 516 AGGCTTATGCGCCAGGCGCAGAGCCAGCAAGCCCTTACTCTCAAGAGAGCTTCTG 575  
 Qy 181 LysTTP\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGlyGluAlaGluLysLeuArgAsp 200  
 Db 576 AATGACCAAGACGACCATCTCATGATGATGCTCGGAGAGGCTGAGAGATCAAGAGC 635  
 Qy 201 IleAlaArgGluValLeuLysLysLys\*\*\*GlyGluTyrSerLeuThrAspPheLys\*\*\*Pro 220  
 Db 636 ATGCTGCGCAGAGTCTTAAAGATCTTGGCGAATACAGCTCTCACTGACTTCACTGCGCA 695  
 Qy 221 LeuLys\*\*\*LeuLysValGlyLysTyrGlyLysArgGlyLeuAspAspLysLeuAsnLysPhe 240  
 Db 696 TTGAAGCATCTCAAGGTTGGAAGATGAGAGAGATGACAGCATCTTGAACAAAGTTT 755  
 Qy 241 AspProValValGluArgValLysLysLysArgArg\*\*\*IleValArgArgArg\*\*\*Asn 260  
 Db 756 GACCTGTGCTTGAAGGCTCATCAAGAGCCGCTGAGATGCTGAGAGAGAGAGAGAAC 815  
 Qy 261 GlyGlu\*\*\*\*\*GluGlyGlu\*\*\*SerGlyVal\*\*\*LeuAspThrLeuLeuGluPheAla 280  
 Db 816 CGAGAGGTTGTGAGGCTGAGGCTGACGCGGCTTCTCTGACACTTCTGAAATTCGCT 875  
 Qy 281 GluAspGluThr\*\*\*GluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 300  
 Db 876 GAGATGAGACATGAGATCAAAATCAACAGACCAAGCAAGCATGCTGTGTGCGAC 935  
 Qy 301 \*\*\*PheSerAlaGly\*\*\*AspSerThrAla\*\*\*\*\*ThrGluThrAlaLeuAlaGluLeu 320  
 Db 936 TTTTCTCGCAGAGACGACTCAAGCGCTGCGAACAGAGTGGCGATTTGGCGAAGCTC 995

Qy 321 IleAsnAspPro\*\*\*ValLeu\*\*\*\*\*AlaArgGluLys\*\*\*TyrSerValAlaGlyLys 340  
 Db 996 ATCAACATCTCTTAAGGTGTGTGAAGAGGCTGTGAAGAGGCTCTCAAGGTGTGTGAAG 1055  
 Qy 341 Asp\*\*\*LeuValAspGluValAspThrGlnAsnLeuProTyrIleArgAlaIleValLys 360  
 Db 1056 GACAGCTGTGACGAGATGACATCAAAACCTTCTTCACTTCAAGAGATGAGAG 1115  
 Qy 361 GluThrPheArgMetHisAspProLeuProValValLysArgLysCys\*\*\*GluLysCys 380  
 Db 1116 GAGACATTTCCGATCGACCCGCCCTCCAGTGTGCAAAAGAAATGACAGAAAGTGT 1175  
 Qy 381 \*\*\*IleAsnGly\*\*\*Val\*\*\*ProGluGlyAlaLeu\*\*\*\*\*PheAsnValIleProGlnVal 400  
 Db 1176 GAGATTAATGATATGATATGATCCAGAGAGGATTTGCTCTCAATGATATGAGCAAGTA 1235  
 Qy 401 Gly\*\*\*Asp\*\*\*LysTyrTTPAspArgProSerGlu\*\*\*ArgProGluArgPheLeuGlu 420  
 Db 1236 GGAGAGACCCCAAAATACGTGGACAGACCATGAGATTCGCTCGAGAGGTTCCTAGAG 1295  
 Qy 421 Thr\*\*\*AlaGluGlyGluAla\*\*\*\*\*LeuAspLeuArgGly\*\*\*HisPheGlnLeuLeu 440  
 Db 1296 ACAGGAGCTGAAGAGGAGAGAGAGGCTCTTGAATCTTGAAGAGCAACATTTTCAACTTCTC 1355  
 Qy 441 PropheGlySerGlyArg\*\*\*MetCysProGlyVal\*\*\*LeuAlaThrSerGly\*\*\*Ala 460  
 Db 1356 CCATTTGAGTCTGGAGAGAGATGTGCTCTGAGATGATCTGCTACTCTGAGAGATGCGA 1415  
 Qy 461 ThrLeuLeuAlaSerLeuLysGlnCysPheAspLeuGlnValLeuGlyProGlnGlyLys 480  
 Db 1416 ACACCTTCTGATCTTATTAAGTCTTCAAGTCTTCAAGTCTGAGAGTCTGAGAGAGAGCAG 1475  
 Qy 481 IleLeuLysGly\*\*\*AspAlaLysValSerMetGluLysArgAlaGlyLeuThrValPro 500  
 Db 1476 ATATTGAAGGCTGTGACGCCAAAGTTACATGACAGAGAGAGCGGCTTCACTGTTCCA 1535  
 Qy 501 ArgAlaHisSerLeuValCysValProLeuAlaArgGlyValAlaSerLysLeuLeu 520  
 Db 1536 AGCGACATGATCTTCTGTCTGTCTTCCACTTGCAGATGCGGCTTGCATTTAACTCTT 1595  
 Qy 521 Ser 521  
 Db 1596 TCT 1598

RESULT 10  
 AF022462  
 LOCUS  
 DEFINITION  
 complete cds.  
 ACCESSION  
 AF022462  
 VERSION  
 AF022462.1  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max.  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;











QY 299 ----- 299

Db 1012 TTGACATTATATATCATGTGGTGCAGTAATTAACGGTACGATCTTAATTCATAT

QY 300 ----- Asp\*\*\*PheSerAlaGly\*\*\*AspSerThrAla\*\*\*\*\*ThrGluTr 315

Db 1072 TGTGATATGACAGACACTTTTCTCGGACAGAACAGACTCCACAGCGGTGGACACAGATG 1111

QY 315 pAlaLeuAlaGluLeuIleAsnAsnPro\*\*\*ValLeu\*\*\*\*\*AlaArgGluGlu\*\*\*Ty 335

Db 1132 GCGATTGGACAGACTCATCAACATCTTAAGGTGTGAAAGGCTCGTAGAGAGGTCTA 1191

QY 335 rSerValValGlyIleAsp\*\*\*LeuValAspGluValAspThrGlnAsnLeuProTyr11 355

Db 1192 CAGTGTGTGGGAAAGACAGACTGTGACAGAGTTGACATCAAACTCTTCAAT 1251

QY 355 eArgAlaIleValIleGluThrPheArgMetHisProProLeuProValValIleArgGly 375

Db 1252 TAGACCAATCGTACAGAGACATTCGCAATGCAACCCGCACTCCAGTGTCAAAAGAA 1311

QY 375 sCys\*\*\*GluGluCys\*\*\*IleAsnGly\*\*\*Val\*\*\*ProGluGlyAlaLeu\*\*\*\*\*Ph 395

Db 1312 GTGCACAGAAAGTGTGAGATTATGATATGATTCACAGAGAGCATTTGATCTCTT 1371

QY 395 eAsnValTrpGlnValGly\*\*\*Asp\*\*\*LysTyrTrpAspArgProSerGlu\*\*\*ArgPr 415

Db 1372 CATGTATGGCAAGTGTGAGAGAGACCCCAATATCTGGGACAGACATCGAGTTCCGTCC 1431

QY 415 oGluArgPheLeuGluThr\*\*\*AlaGluGlyGluAla\*\*\*\*\*LeuAspLeuArgGly\*\* 435

Db 1432 TTAGAGGTTCTCTGAGACACAGGCGCTGACAGGAGAGAGGCGCTTGTATCTAGGGGACA 1491

QY 435 \*HisPheGluLeuLeuProPheGlySerGlyArg\*\*\*MetCysProGlyVal\*\*\*LeuAl 455

Db 1492 ACATTTTCAACTTCTCCATTTGAGTCTGAGAGAGATGTGCTCTGAGTCAATCTGAC 1551

QY 455 aThrSerGly\*\*\*AlaThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnValIle 475

Db 1552 TACTTCGGGAATGGACACACTTCTGCACTCTTAATTCAGTCTTCCACTTCCAGAGTCTT 1611

QY 475 uGlyProGlnGlyGlnIleLeuLysGly\*\*\*AspAlaLysValSerMetGluGluArgAl 495

Db 1612 GCGTCCAAAGACAGATATTGAAGGAGTGTGACGCCAAAGTTACATGACAGAGAGAC 1671

QY 495 aGlyLeuThrValProArgAlaHisSerLeuValCysValProLeuAlaArgGlyIleGly 515

Db 1672 CCGCTCACTGTCCAAAGGACATAGTCTGTGTGTGATTCACATTCGAAGAGATCGGCGT 1731

QY 515 lAlaSerLysLeuLeuSer 521

Db 1732 TGCATCTAAACTCTTCTT 1750

RESULT 13

LOCUS AF195818 1800 bp DNA linear PLN 23-MAR-2000

DEFINITION Glycine max isoflavone synthase 1 (ifsl) gene, partial cds.

ACCESSION AF195818

VERSION AF195818.1 GI:7288452

KEYWORDS Glycine max.

SOURCE Glycine max

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 1800)

AUTHORS Jung, W., Yu, O., Lau, S.M., O'Keefe, D.P., Odell, J., Fader, G. and McConigle, B.

TITLE Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes

JOURNAL Nat. Biotechnol. 18 (2), 208-212 (2000)

MEDLINE 20124255

PUBMED 10657130

REFERENCE 2 (bases 1 to 1800)

AUTHORS Jung, W., Yu, O., Odell, J., Fader, G. and McConigle, B.

TITLE Direct Submission

JOURNAL Submitted (18-OCT-1999) Nutrition and Health, DuPont, PO Box 80402, Wilmington, DE 19880-0402, USA

FEATURES

source location/Qualifiers

1..1800

/organism="Glycine max"

/db\_xref="taxon:3847"

1..1800

/gene="ifsl"

join<1..893,1112..>1800

/gene="ifsl"

/product="isoflavone synthase 1"

join<1..893,1112..>1800

/gene="ifsl"

/note="cytochrome P450"

/codon\_start=3

/product="isoflavone synthase 1"

/protein\_id="AAF45142.1"

/db\_xref="GI:7288453"

/translation="LEIALGLFVALFLIARPTSAKSKALHLNPSPSPRLPRIG  
HHHLIKRLHVALIDLSKKHGPFSLSGSMPTVAYSTELFKFLDTQIATRSNTR  
FQTSAIRRLTYDNSVAMVFGPYKVKFVKRLIMNDLNAVTVNKLPLRTOIRKFLRV  
MQSAEIQKELDVTEBELKNTNSTISMMLGAESEIRDLARVLKIPGEVSLDPTLP  
LKYLVKYEKRIIDLINKPDPVVERIVKREIEVRRKNGEVEGSAAGVFDLTLL  
FAEDTEIEIKITKEQIKGLVVDPSAGTSTAVATEMALAEALINNRPVLOKAREEVS  
VIGDRLVDEVDTQNLPIYRAIYKETFRRHPPLVIVRKCTECCINQYVPEGALVL  
FVWVGAGDVKYMDRPSSEFRPERLETGAEAGLDLRQGHQLLPFGSGRMKCPGV  
NLAISGATLLASLIQCFLDQVLQPOGQILKGDADAKVSMERAGLTVRAHSLVCVPL  
ARIGVASKSLLS"

BASE COUNT 476 a 441 c 413 g 470 t

ORIGIN

Alignment Scores:

Pred. No.: 3.04e-250 Length: 1800

Score: 2189.00 Matches: 451

Percent Similarity: 76.18% Conservative: 0

Best Local Similarity: 76.18% Mismatches: 68

Query Match: 91.36% Indels: 74

DB:	8	Gaps:	1
US-09-857-561-66 (1-521) x AF195618 (1-1800)			
Qy	3 LeuGIuLeuAlaLeuGIyLeu***ValLeuAlaLeuPhe***HisLeuArgProThrPro	22	
Db	3 CTGGAAGCTGCACTGGTTGTTGTTAGCTTTGTTTGTGACTGCTGCGCCACACA	62	
Qy	23 ***Ala***SerIyAlaLeuArgHisLeuProAsnProSerPro***ProArgLeu	42	
Db	63 AGTGAATATCAAAAGCACTGCGCACCTCCAAACCTCCAAAGCAAGCTGCTTT	122	
Qy	43 ProPheIleGIyHis***HisLeuLeuAspIyLeuLeuHisTyrAla***IleAsp	62	
Db	123 CCTTCATGTGGCCACTTCCTCTTAAAGATTAACCTTCCTCACTAGCACTCATCGAT	182	
Qy	63 LeuSerIyLysHisGIyProLeuPheSer*****PheGIySerMetProThrValAla	82	
Db	183 CTCTCAAAAAGCATGGCCCTTAATCTCTCTCTCTGCTGCTGCTCATCCCAACCGTGT	242	
Qy	83 AlaSerThrProGIuLeuPheIyLeuPheLeuGIu***GIuAlaThrSerPhe***	102	
Db	243 GCTTCACCCCTGAGTGTTCAGCTCTTCCTCAACCCAGAGGCAACTTCCTTCAC	302	
Qy	103 ThrArgPheGIuThrSerAla***Arg***LeuThrTyrAsp***ValAla*****	122	
Db	303 ACAAGTTCCAAACTCTGCATTAAGCCCTCACTTAAGCACTGCTGGCCATGTGTT	362	
Qy	123 Pro***GIyProTyrTrp***PheValArgIyLeuIleMetAsnAspLeuAsnAla	142	
Db	363 CCATTGCGACCTTACTGGAAGTCTGTGAGCAAGCTCATGATCAAGCACTTCACACGCC	422	
Qy	143 ThrThrValAsn***LeuArgProLeuArgThrGIuGIuIleArgIyLys***LeuArg***	162	
Db	423 ACACACGTCACAAAGCTCAGGCTTTGAGAGACCAAGATCCGCAAGTTCCTTAAGGTT	482	
Qy	163 MetAlaGIu***AlaGIuAla***LysProLeuAsp***ThrGIuGIuLeuLeuLysTrp	182	
Db	483 ATGGCCCAAGCCGAGAGGCCGAGAGCCCTTGAACCTCAAGAGACTTCCTCAATGG	542	
Qy	183 ***AsnSerThr***SerMetMet***LeuGIyGIuAlaGIuGIuIleArgAspIleAla	202	
Db	543 ACCAAAGACACATCTCATGATGATGCTCGGCGAGGCTGAGAGATCAGAGACATCGCT	602	
Qy	203 ArgGIuValLeuLysIle***GIyGIuTyrSerLeuThrAspPheIle***ProLeuLys	222	
Db	603 CCGAGGATCTTAAGATCTTCGCGAGATCAAGCCCTCACTCATCTGCGCTTTGAAG	662	
Qy	223 ***LeuLysValGIyLysTyrGIuLysArgIleAspAspIleLeuAsnLysPheAspPro	242	
Db	663 TATCTCAAGGTGGAAAGTATGAAAGAGATGATGATCATCTTGAACAGTTTCAGCCT	722	
Qy	243 ValValGIuArgValIleLysLysArgArg***IleValArgArgArg***AsnGIyGIu	262	
Db	723 GTCTTAAAGGCTCATCAAGAGCCCTCATGATCTCTCAAGAGAGAAAGAGAGAGAA	782	
Qy	263 *****GIuGIyGIuLys***SerGIyVal***LeuAspThrLeuLeuGIuPheAlaGIuAsp	282	
Db	783 GTTGTGAGGCGAGAGCGGCGCTCTCTCGACACTTCCTGATTCGCTGAGAGAC	842	
Qy	283 GIuThr***GIuIleLysIleThrLys***IleLysGIyLeuValVal-----	299	
Db	843 GAGACATGAGATGAAATTTTCCAGAGAGCAAAATCAAGAGGCTGTGTGT-CGTAAAGTTT	901	
Qy	299 -----	299	
Db	902 CCTTCTCTCTCTACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	961	
Qy	299 -----	299	
Db	962 TATACTATATGAGAAATATATATGACGACTACGAGTAAAGATATGTGTGTTTTTTA	1021	
Qy	299 -----	299	
Db	1022 AAAAGATACAAAGTCTTTATGATGATGATGATGATGATGATGATGATGATGATGATG	1081	
Qy	300 -----	309	
Db	1082 ACTAATTAATTCATATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATG	1141	
Qy	310 Ala*****ThrGIuThrPalalaLeuAlaGIuLeuIleAsnAsnPro***ValLeu*****	329	
Db	1142 GCGGTGCAACAGAGAGGCGCATGCGCAGAGCTCAACATCCAGAGGTGTTCACAAAG	1201	
Qy	330 AlaArgGIuGIu***TyrSerValValGIyLysAsp***LeuValAspGIuValAspThr	349	
Db	1202 GCTCTGAGAGAGGTCTACAGAGTGTGTGGCAAGATGATCTGTGAGAGAGTTCACCT	1261	
Qy	350 GIuAsnLeuProTyrIleArgAlaIleValLysGIuThrPheArgMetHisProProLeu	369	
Db	1262 CAAAACCTCTCTTACATTAGGCGCATTTGAGAGAGATTCGGAATGACACCACTC	1321	
Qy	370 ProValIyLysArgLysCys***GIuGIuLys***IleAsnGIy***Val***ProGIu	389	
Db	1322 CCGTGTCAAAAGAAAGTGCAGAAAGTGTGATTAATGCGATGATGATGATGATGATG	1381	
Qy	390 GIyAlaLeu*****PheAsnValITPginValGIy***Asp***LysTyrITPAspArg	409	
Db	1382 GGAGCATTTGCTCTTTCAATGTTTGGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG	1441	
Qy	410 ProSerGIu***AspProGIuArgPheLeuGIuThr***AlaGIuGIyGIuAla*****	429	
Db	1442 CATCAGAAATTCGTCGAGAGAGTCTTGAAGACTGCTGTAAGAGAGAGAGAGAGAG	1501	
Qy	430 LeuAspLeuArgGIy***HisPheGIuLeuLeuProPheGIySerGIyArg***MetCys	449	
Db	1502 CTTCATCTTAAGGCGCAGCATTTCCAACTCTCCCATTTGGCTTGGAGAGAGATGTGC	1561	
Qy	450 ProGIyVal***LeuAlaThrSerGIy***AlaThrLeuLeuAlaSerLeuIleGIuLys	469	
Db	1562 CCGTGTGCAATTTGCTTCACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1621	
Qy	470 PheAspLeuGIuValLeuGIyProGIuGIuGIuIleLeuLysGIy***AspAlaLysVal	489	



AF195800	1501 bp	mrna	linear	PLN 16-FEB-2000
LOCUS				
DEFINITION	Medicago sativa isoflavone synthase 1 (ifsa1) mRNA, partial cds.			
ACCESSION	AF195800			
VERSION	AF195800.1	GI:6979523		
KEYWORDS				
SOURCE	Medicago sativa.			
ORGANISM	Eukaryota; Vitisidplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustroids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.			
REFERENCE	1 (bases 1 to 1501)			
AUTHORS	Jung, W., Yu, O., Lau, S.-M., O'Keefe, D. P., Odeh, J., Fader, G. and McGonigle, B.			
TITLE	Identification and expression of isoflavone synthase, the key enzyme for biosynthesis of isoflavones in legumes			
JOURNAL	Nat. Biotechnol. 18 (2), 208-212 (2000)			
MEDLINE	20124255			
PUBMED	10657130			
REFERENCE	2 (bases 1 to 1501)			
AUTHORS	Jung, W., Yu, O., Odeh, J., Fader, G. and McGonigle, B.			
TITLE	Direct Submission			
JOURNAL	Submitted (18-Oct-1999) Nutrition and Health, DuPont, P. O. Box 80402, Wilmington, DE 19880-0402, USA			
FEATURES	location/Qualifiers			
SOURCE	1..1501			
gene	/organism="Medicago sativa"			
	/db_xref="taxon:3879"			
	<1..>1501			
	/gene="ifsa1"			
	<1..>1501			
	/gene="ifsa1"			
	/codon_start=3			
	/product="isoflavone synthase 1"			
	/protein_id="AA034521.1"			
	/db_xref="GI:6979524"			
	/translation="PFLHRLPESKASKLSEHLNPPSPKRLPFIQHLHLKDLKLLHLYALIDLSKHKQPLPESLPGSMPTVASTPELFLKLTQHEATSPNTRPOTASPRLLTYINSAWVPGGYMFPRLKILKMDLLATTVYKRLPRLQIRKLRYVMQASBAKQPLDVELELTKNTSTISMMMLGAEIEIKDINRELVKIFGEYSLTDPWPLVYKLVKTEKRSIDILINKDDEVERIVIKRGGIVRRENGVEVGCASGVPLDTLLEPMEDEMERIKI1KEQIGALVVDLFSAGDSTAVATMALNLINPNVLQKAREEVSVVGDRLVDEVKITQNLVYIALVKEITRMHPPLPVYKRRKCTEECINGVYIEGALVLFVWQVGRPKVYKMPSEPPRRRLTGAAGAGAPLDLRGCHOLLPGSGRMCPGNVILATSGMALLLRLTQCDLQVAGCGQILKGDPAKSMERAGLTPRAHSVVCPLARIG"			
BASE COUNT	390 a 400 c 366 g 345 t			
ORIGIN				
Alignment Scores:				
Fred. No.:	2,86e-249	Length:	1501	
Score:	2180.00	Matches:	433	
Percent Similarity:	86.77%	Conservative:	0	
Best Local Similarity:	86.77%	Mismatches:	66	
Query Match:	90.98%	Indels:	0	
DB:	8	Gaps:	0	

Qy 15 Phe\*\*\*HisLeuArgProThrPro\*\*\*Ala\*\*\*SerIleAlaLeuAlaGhiIleuProAsn 34  
 Db 3 TTCTGCACTGCTGCTCCACCAAGTGCMAAATCAAGCACTTCGCACTCCCAAC 62  
 Qy 35 ProProSerPro\*\*\*ProArgLeuProPheIleGlyHis\*\*\*HisLeuLeuValAspLys 54  
 Db 63 CCCCCAGGCCCAAGCTGCTCTCCCTTCATGAGCACTTCCTTAAAGATATA 122  
 Qy 55 LeuLeuHisIleVal\*\*\*IleAspLeuSerLysHisGlyProLeuPheSer\*\*\*\*\* 74  
 Db 123 CTTCGCACTATGCACTCATGATCTCTCAAAAAGCATGGCCCTTATCTCTCTCC 182  
 Qy 75 PheGlySerMetProThrValValAlaSerThrProGluLeuPheLysLeuPheLeuGln 94  
 Db 183 TTCGGCTCCATGTCACCGCTGTTGCTCCACCCCTGATGTTGTCAGCTTCCTCCAA 242  
 Qy 95 \*\*\*\*\*GluAlaThrSerPhe\*\*\*ThrArgPheGlnThrSerAla\*\*\*Arg\*\*\*LeuThr 114  
 Db 243 ACCCAAGAGCAACTTCCTTCACCAAGGTTCCAACTTCGCAAGAGAGCCCTCACT 302  
 Qy 115 TyrAsp\*\*\*\*\*ValAla\*\*\*\*\*Pro\*\*\*GlyProTyrTyr\*\*\*PheValArgLysLeu 134  
 Db 303 TACCACACTCTGTGGCCATGCTTCATTCGACCTTACGAGGTTGTGAGAGAGCTC 362  
 Qy 135 IleMetAsnAspLeuLeuAsnAlaThrThrValAsn\*\*\*LeuArgProLeuArgThrGln 154  
 Db 363 ATCATGAAGCACTTCCTCAAGCCACCAACCGTCAACAGCTGAGCCCTTGAAGACCAA 422  
 Qy 155 GlnIleArgLys\*\*\*LeuArg\*\*\*MetAlaGln\*\*\*AlaGluAla\*\*\*LysProLeuAsp 174  
 Db 423 CAGATCCGCAAGTCTTAAAGGTTATGGCCCAAGCCCAAGGCCCAAGGCCCTTGAC 482  
 Qy 175 \*\*\*ThrGluGluLeuLeuLysTyr\*\*\*AsnSerThr\*\*\*SerMetMet\*\*\*LeuGlyGlu 194  
 Db 483 GTCACGAGAGAGCTTCTCAATAGAACCAAGCAACATCTTCATGATGATGCTGCGCAG 542  
 Qy 195 AlaGluGluIleArgAspIleAlaArgGluValLeuLysIle\*\*\*GlyLysTyrSerLeu 214  
 Db 543 GCTAGAGATGCAAGACATCCCTGCGCGAGTCTTAAAGTCTTGGCGCATACAGCCCTC 602  
 Qy 215 ThrAspPheIle\*\*\*ProLeuLys\*\*\*LeuLysValGlyLysTyrGlyLysArgIleAsp 234  
 Db 603 ACTGACTTATCTGGCCTTGAAGTATCTCAAGGTTGGAAGTATGGAAGAGAGATGAT 662  
 Qy 235 AspIleLeuAsnLysPheAspProValValGluArgValIleLysLysArgArg\*\*\*Ile 254  
 Db 663 GACATCTTGAACAGATTCGACCTGCTGAAGGGTCAATCAAGAGCCCGTGGGATC 722  
 Qy 255 ValArgArgArg\*\*\*AsnGlyGlu\*\*\*\*\*GluGlyGlu\*\*\*SerGlyVal\*\*\*LeuAsp 274  
 Db 723 GTCAGAAAG 782  
 Qy 275 ThrLeuLeuGluPheAlaGluAspGluThr\*\*\*GluIleLysIleMetLys\*\*\*\*\*Ile 294  
 Db 783 ACTTGTCTGATTCGCTGAT 842

Qy 295 LysGlyLeuValValAlaAsp\*\*\*PheSerAlaGly\*\*\*AspSerThrAla\*\*\*\*\*ThrGlu 314  
 Db 843 AAGGACCTTGTGTTGCACTTCTCTGCAAGGACAGATTCACAGCGGTGGCAACAGAG 902  
 Qy 315 TrpAlaLeuAlaGluLeuIleAsnAsnPro\*\*\*ValLeu\*\*\*\*\*AlaArgGluGlu\*\*\* 334  
 Db 903 TGGGATTTGGCAGAGCTCATACATCCAGGAGGTTGGCAAAAGGCTGGAGAGGCTC 962  
 Qy 335 TyrSerValValGlyLysAsp\*\*\*LeuValAspGluValAlaAspThrGlnAsnLeuProTyr 354  
 Db 963 TACAGTGTGTGGCAAGAGATAGACTGTTAGACAGATGTACACTCAAACTTCCTTAC 1022  
 Qy 355 IleArgAlaIleValLysGluThrPheArgMetHisProProLeuProValValLysArg 374  
 Db 1023 ATTAGGGCACTTGTGAAGAGACATTCGATGACACCACTCCAGTGTGCAAAAGA 1082  
 Qy 375 LysCys\*\*\*GluGluLys\*\*\*IleAsnGly\*\*\*Val\*\*\*ProGluGlyAlaLeu\*\*\*\*\* 394  
 Db 1083 AAGTCAACAGAGAGAGTGAAGATTAATGGATATGTATGCCAGAGAGAGCATTTGCTT 1142  
 Qy 395 PheAsnValTrpGluValAlaGly\*\*\*Asp\*\*\*LysTyrTrpAspArgProSerGlu\*\*\*Arg 414  
 Db 1143 TTCAATGTTGGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1202  
 Qy 415 ProGluArgPheLeuGluLysThr\*\*\*AlaGluGlyGluAla\*\*\*\*\*LeuAspLeuArgGly 434  
 Db 1203 CCGAGAGAGTCTTGAAGAACTGTGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1262  
 Qy 435 \*\*\*HisPheGlnLeuLeuProPheGlySerGlyArg\*\*\*MetCysProGlyVal\*\*\*Leu 454  
 Db 1263 CAGCATTTCCAACTCTCCCTGCTGGGTGAGAGAGAGATGCTGCTGCTCAATTTG 1322  
 Qy 455 AlaThrSerGly\*\*\*AlaThrLeuLeuAlaSerLeuIleGlnCysPheAspLeuGlnVal 474  
 Db 1323 GCTACTTCAGAGATGGAACACTTCTTGATCTCTTATCCATAGCTTTGACCTGCAAG 1382  
 Qy 475 LeuGlyProGluGlnIleLeuLysGly\*\*\*AspAlaLysValSerMetGluGluArg 494  
 Db 1383 CTGGGCTCTCAAGAGCAATTTCAAGAGTATGATGATGATGATGATGATGATGATGAT 1442  
 Qy 495 AlaGlyLeuThrValProArgAlaHisSerLeuValCysValProLeuAlaArgIle 513  
 Db 1443 GCTGGCTTCACAGTTCGAAGGAGCAATAGTCTGTTGTGTTTCCACTTCAGAGATC 1499

Search completed: May 18, 2003, 13:07:12  
 Job time : 2876 secs

GenCore version 5.1.4\_p5\_4578  
 Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
 Run on: May 18, 2003, 07:48:03 : Search time 239 Seconds  
 (without alignments)